Draft: A Distributed Science Cybersecurity Program

(Please send any comments or questions regarding this draft white paper to DAAgarwal@lbl.gov.)

Abstract

The Department of Energy (DOE) Office of Science and the National Science Foundation (NSF) are responsible for the operation of some of the nation's most advanced research and development user facilities located at the national laboratories and universities. These state-of-the-art facilities are shared with the science community worldwide, and contain technologies and instrumentation that are available nowhere else. They include particle and nuclear physics accelerators, synchrotron light sources, neutron scattering facilities, supercomputers, and high-speed computer networks. Each year, the DOE Office of Science facilities are used by more than 18,000 researchers from universities, other government agencies, and private industry. NSF supercomputer centers also provide compute resources to researchers around the world. Providing open access to these resources presents an enormous security challenge. A compromised system at any researchers' institution must not directly lead to a security compromise at the corresponding user facility.

DOE Office of Science and NSF researchers are also participants in experiments such as ITER, CMS, and ATLAS, which are hosted by other countries. These one-of-a-kind experiments involve thousands of scientists spread throughout the globe, including sensitive countries. The NSF Network for Earthquake Engineering Simulation (NEESGrid) and National Virtual Observatory (NVO) projects are providing users with access to resources around the country. These DOE and NSF user facilities and research collaborations can ill afford to be offline for extended periods due to security incidents.

Scientific discovery has become a global, collaborative, distributed endeavor involving access to a broad range of computer resources to complete an experiment, run a simulation, or search databases. There are several projects building the applications and middleware to implement the infrastructure needed by distributed science; however, little attention has been paid to the cybersecurity of these environments. A program to develop the next generation of cybersecurity solutions to support user facilities and distributed science needs to be undertaken. This program will bridge the ever-widening gap between cybersecurity research and the cybersecurity tools that are in use today. Success of the program will depend on collaboration between researchers, stakeholders, policy makers, users, and site cybersecurity personnel to develop and deploy the cybersecurity mechanisms needed to protect science resources and participation in distributed science projects into the future.

1. Introduction

The DOE Office and Science and the NSF support many user facilities. For example, the National Synchrotron Light Source at Brookhaven National Laboratory is the world's brightest continuous source of X-rays and ultraviolet radiation for research. The Environmental Molecular Sciences Laboratory at Pacific Northwest National Laboratory houses one of the world's most powerful widebore nuclear magnetic resonance (NMR)

spectrometers. The Spallation Neutron Source, being built at Oak Ridge National Laboratory, will provide the most intense pulsed neutron beams in the world for scientific research and industrial development. Other DOE user facilities include the Advanced Light Source, the Advanced Photon Source, the National Energy Research Scientific Computing Center (NERSC), and the new National Leadership Computing Facility. These facilities are sited at DOE laboratories and are utilized by a broad range of users from across the country and the world. NSF facilities include the San Diego Supercomputer Center, the National Center for Supercomputing Applications, the National Center for Atmospheric Research, and the National Laboratory for Applied Network Research. Biology projects such as the Shewanella Federation are bringing together different studies of the same bacteria to develop a system-level understanding. Shared massive databases supporting both national and international collaborations are also becoming common. Some collaborations include massive simulations to study global problems such as global warming. An important commonality between these facilities and projects is the need for high performance computational and networking resources, and a distributed collection of researchers needing to use them.

Although many of the major science experiments that DOE and NSF researchers are participating in are located at DOE and NSF facilities, increasingly, key experiments such as CMS and ATLAS are located elsewhere. The CMS and ATLAS experiments at the Large Hadron Collider (LHC) being built in CERN, Switzerland each involve approximately 2000 physicists from around the world. DOE, as the host of the U.S. CMS and ATLAS Tier 1 centers, is providing a key element of the global support infrastructure for these experiments. NSF is also providing key infrastructure to these experiments by sponsoring universities to host the Tier 2 centers. The ITER fusion reactor will be located in France and will involve fusion scientists from all over the world in its operations and experiments. In these experiments, the software used for computing and access control is defined through international agreements.

These distributed collaborative projects typically form a *virtual organization*. Within these virtual organizations, the authentication and authorization are federated to enable cross-site authentication, incorporate dynamically available resources, and manage allocation of resources. This virtual organization model redefines the traditional enclave into one that crosses and incorporates many individual site borders and includes personnel and resources from a wide range of sites. Within the confines of the virtual organization, large quantities of data need to be able to move using high-speed communication links from site to site.

At the same time that these new distributed collaborative science models are developing, new, more malicious cybersecurity attacks are also happening. The recent attack on several NSF and DOE supercomputing centers is an example of this, including an attack that took the San Diego Supercomputer Center off the network for an entire week. DOE and NSF sites, as key participants in these collaborations, need methods of participating as first class entities in this open science model while protecting their resources from hackers. This virtual organization will mean that users from across the collaboration will have access to DOE and NSF resources. However, an attack and possible compromise at a single site within the virtual organization must not directly compromise other sites and resources. Protections of high performance computers are also particularly important

because recovery of a compromised supercomputer system can take weeks, during which the machine is unavailable for its users.

As ESnet, UltraScience Network, Internet2, and National LambdaRail move to 40 Gbit connections at high-end computing facilities, dedicated wavelength services will be used to connect supercomputing facilities for specific applications and to transfer data. In order to achieve the desired throughput, these wavelength services will often bypass cybersecurity mechanisms in place for regular network connections, leaving an unprotected and unmonitored path between centers. A method of allowing these dedicated circuits to achieve their purpose without compromising the security of an individual site is needed.

Deployment of each individual component in a cybersecurity system is a balance between cost and benefit. The available choices of tools need to be flexible enough so that decisions regarding which components are needed at a particular enclave can be determined locally, based on risk analysis and an expected return-on-investment (ROI) calculation. Although best practices help inform the decision of what components to deploy, this approach, if used alone, can miss important threats not typically found in common best practices scenarios. Through use of accurate ROI data, decisions can be made based on expected cost of a protection and expected benefit of the protection. Calculating ROI for cybersecurity protections is still a developing science, and further development and refinement can only be accomplished by comparing real performance and cost of a protection mechanism with the expected performance. The dividends of improving this technique are obvious — it has the potential to turn the design of cybersecurity defenses from a black art into a science. Intelligent cybersecurity deployment decisions for distributed science environments are particularly critical since the design space is almost infinite.

Current research in cybersecurity mechanisms is showing a great deal of promise, but significant development is required before this potential can be realized. In addition, the commercial world is working on some of these issues, and some of the new cybersecurity products are quite impressive in terms of performance and functionality. However, the DOE Office of Science and NSF are responsible for the operation of some of the nation's most advanced research and development user facilities. These resources have unique security requirements, including cybersecurity tools that are customizable and have the ability to incorporate log data from specialized resources running specialized operating systems. Commercial off-the-shelf (COTS) security products are not designed to secure these specialized resources. Another issue for these environments is interoperability. In general the COTS solutions have proprietary interfaces and can only interoperate with other products from the same vender.

While COTS tools were developed with businesses in mind, open science is a very different environment — one far more varied and demanding. The cybersecurity solution for the future of DOE Office of Science and NSF research will require a *combination* of existing COTS solutions and research solutions, as well as a new framework to allow security components to exchange information.

The rest of this paper outlines a program of work that could be undertaken by the DOE Office of Science and the NSF to enable them to participate in distributed science collaborations while maintaining open access and cybersecurity at individual sites.

2. Approach

The DOE Office of Science and NSF cybersecurity program focus will be on developing cybersecurity capabilities to protect a distributed, collaborative science environment while maintaining as open an environment as possible. The program needs to address several key areas to accomplish this. These areas include:

- cybersecurity for high performance networking and computing environments
- providing open access while protecting resources
- coordination between cybersecurity components to cooperatively recognize and react to cyber threats
- integration of middleware security mechanisms with the cybersecurity approach, and implementation of one-time password (OTP) and Personal Identity Verification (PIV) mechanisms in the environment.

Key elements of the program will include deployment and testing of both COTS and research cybersecurity solutions in operational environments, and research into the additional tools and components needed in an operational cybersecurity setting. The end goal is to provide practical solutions and tools which can be deployed within research collaborations funded by the NSF and the DOE Office of Science.

High Performance Computing and Networking

High performance computing resources are a centerpiece of the DOE Office of Science program and are also important in NSF. These resources bring with them cybersecurity challenges far different than those in a typical environment, but which must be addressed in order to maintain the appropriate level of both security and availability. High performance computing machines are high-profile resources, which generally run special operating systems and software. Recovery often requires handcrafted procedures which can take weeks to complete after a compromise. Typical cybersecurity practices would lock down such high-value resources and allow only limited access to a few trusted users. However, this approach runs counter to the mission of the DOE Office of Science national user facilities and NSF high performance computing resources. Instead, improved mechanisms need to be developed to prevent, detect, and recover from compromise, in addition to gathering forensics data, while operating at the speeds required in a high performance computing environment.

High performance networking capabilities at speeds of 10–40 Gbits per second are being deployed to connect DOE Office of Science labs and NSF facilities, enabling distributed collaborative science and providing access to high performance computing resources. These links will provide unprecedented capabilities to regularly transport large data sets, and enable collaborative science to the degree that many science programs are designing their computing models to take advantage of these capabilities. It is important that these

high-speed paths be protected without compromising the operational capabilities of the network link.

The new high speed networks make it impossible with today's technology to have a standard firewall and intrusion detection system implementation, as they cannot handle the volume of data being transported. New approaches to securing these networks are necessary. The technology is moving faster than security ever will. Nevertheless, if the U.S. is going to stay competitive with the rest of the world, we are going to have to use these new technologies, whether they are secure or not. Now is the time to get started making sure they are secure.

This program will develop solutions for traffic analysis, intrusion detection, and intrusion prevention on these high-speed links. It will also create mechanisms to quickly detect and recover from intrusions on high performance computing resources.

Connecting Cybersecurity Components

Today a remote user accessing a facility typically passes several cybersecurity barriers to obtain access. First, the user's own local-area network may provide route and or content filtering. In the wide-area network, generally only route filtering and monitoring take place. At the destination border, traffic must pass through the enclave protections such as firewalls and intrusion detection systems. Inside, traffic must pass through the local-area network at the destination enclave. When it reaches the facility host, the traffic will typically be tested again for protocol and port, and might have to pass a firewall or other additional filtering. On the host, the user must then authenticate at the operating system or service level, and pass appropriate authorizations for the attempted resource access. Crypto-card and one-time password can be and often are used at many of these levels to provide strong authentication.

At most enclaves, these layers operate independently and do not actively share information about cybersecurity protections or violations. In addition, enclave cybersecurity, host/resource security, and software security are completely separate functions and the responsibility of separate groups. The enclave cybersecurity personnel protect the enclave from intrusion, while the resource maintainers provide access to legitimate users. Software security then limits the allowed actions of individuals based on fine-grained authorization and protects the data end-to-end by encrypting it.

Although defense-in-depth and diversity of cybersecurity protections are important to protecting an enclave, coordination between the components of the system has the potential to dramatically increase the overall level of security effectiveness and to dramatically reduce the propagation opportunities available to hackers. For example, the authentication and authorization could be coordinated between components to allow enclave border mechanisms to recognize unauthorized connections and activities. With the addition of host-based protections, violations of policy can be configured to restrict site access for the user or place the user on a special watch list for the site. Coordination between components of the cybersecurity system can also enable sharing of recent activity alerts to enable dynamic protection from new or developing threats. Information collection across cybersecurity mechanisms can also help with recognition of and protection against attacks, since many attacks and their extent are only apparent once

information from across one or more enclaves is collected. This information also provides the basis for the response to and recovery from the attack.

This program will develop a framework for coordination between security components and merging information from multiple security layers to provide a more complete view, providing for a more complete and accurate cybersecurity perspective.

Enabling and Protecting Virtual Organizations

Virtual organizations incorporate resources and users from many sites and thus are placed in a unique position when it comes to cybersecurity. These organizations are responsible for controlling resource allocations and for authenticating users. Protection mechanisms are needed that enable virtual organizations to work with sites to minimize risks and enforce policies. Authentication mechanisms such as OTP/PIV, which are beginning to be used by the virtual organization and/or the site, must be integrated into the entire end-to-end system. Each point where an authentication token will be verified or translated to another credential or proxy must be carefully designed and analyzed to ensure that vulnerabilities are not inadvertently introduced. All DOE laboratories are expected to have PIV mechanisms in place by October 2006, so deploying, integrating, testing, and verifying the security of these mechanisms quickly is critical.

The end-to-end system includes middleware and application mechanisms, specialty resources, and site access mechanisms common to both the end site and the virtual organizations. In addition, large cross-agency and global collaborations bring with them their own software and operating system requirements (e.g., LHC Grid Computing). This software is often integral to participation in the collaboration, and may require specific supercomputing operating systems to be installed or changed. Improved vulnerability containment capabilities are needed to allow DOE Office of Science and NSF researchers to participate in these collaborations without risks to the infrastructure.

This program will develop tools and services to allow virtual organizations to better monitor their resources and perform incident containment.

Analysis/Response/Forensics/Recovery

Incident response typically needs to be coordinated between the local resource, local network, border, virtual organization, and wide-area network. This coordination needs to be automated to the extent possible, since hackers rarely attack during normal business hours and often exploit the system(s) and move on within minutes or hours of the initial attack. There are two keys to effective incident response. The first is accurate information and analysis of the attack. The more information the responder has about the scope and type of attack, the more effectively they can respond. The second key is communication. Typically a response requires coordination between several individuals, possibly at different sites, within a small time window. This is a critical problem for incident response and handling.

Forensic data from all levels of the system are critical to long-term response (e.g., prosecution) and effective recovery. There are two primary goals: to collect evidence and to minimize recovery effort. This data is often high volume and from a diverse set of sources. The responders need to be able to analyze the data and determine exactly which hosts have been compromised and the nature of the compromises in order to contain the

attack and narrow the recovery effort. One of the issues that makes incident response difficult is that each collaborating site has different data available. Each site will always use different software, log at different levels, log using different criteria, and so on. Mechanisms are needed to collect and normalize all this data in a very short time frame.

A related service needed by system administrators and security staff responsible for responding to an incident is a secure and authenticated communication channel. The team of responders must be able to rapidly create a communication channel for responding to incidents, using a suite of secure information and communication services tuned to the needs of security officers and their partners.

The ability to quickly recover from a security incident adds the additional value of allowing fast recovery from non-malicious user errors. In fact, user or administrator errors can cause as much damage as a malicious hacker. It is also important to be able to quickly determine when problems are due to unauthorized activities and when they are due to activities triggered by legitimate members of the user communities.

Once a hacker has compromised a system, recovery needs to be rapid, efficient, and thorough to minimize cost and latent risk as much as possible. If the hacker obtained root privileges, then all file systems and files (including the operating system) must be considered suspect. Any file in the system might be modified (e.g., ssh with an embedded keyboard sniffer) or corrupted. Mechanisms to quickly determine what if anything has changed on the system are essential to rapid recovery. The current approach of reinstalling the system and all software from scratch is time consuming and expensive and does not address the issues of possible corruption of user data files.

This program will develop tools for configuration verification and secure authenticated communication, as well as tools for aiding incident response. These are essential and often overlooked components to an overall security architecture.

3. Teaming of Researchers and Operators

A critical element of the proposed program will be the extensive teaming of operations staff, system administrators, stakeholders, policy makers, and end-users along with the cybersecurity researchers in the development and deployment of security related systems. The needs and issues of each of these groups must be confronted and addressed during the design and deployment of the system. In particular, the program must provide explicit funding for and incentives for direct involvement of the production system administrators and cybersecurity personnel. Such a partnership will allow COTS and research cybersecurity mechanisms to be run and tested in production networks and systems.

The Department of Energy national laboratories and large NSF centers provide an ideal environment in which to build, test, and deploy a cooperative cybersecurity system. There is a unique level of consensus and information interchange across and within enclaves — collaboration grown from the fact that the labs and their wide-area network are all under the Department of Energy. It is also aided by the fact that many of the science programs are based on large collaborations that span many enclaves and cross national borders. In addition, some projects such as the Open Science Grid span both DOE and NSF facilities and networks. These provide an ideal environment for deploying

and testing interoperable cybersecurity systems across enclaves with different levels of trust.

4. Interaction with Other Programs

The NSF Cybertrust program is funding 33 projects which are primarily focused on traditional cybersecurity research areas and development of prototypes tested on the EMIST testbed. Although the CyberTrust program is pursuing valuable research, very little of it will have significant impact on the cybersecurity issues facing distributed science. We expect that the NSF CyberTrust program will provide some of the tools needed for the new program and hopefully will work with the new program to tackle some of the cybersecurity issues faced by distributed science in future projects. The NSF Center for Internet Epidemiology and Defenses, the STIM Center (Security Throughput Interaction Modeling), and Team for Research in Ubiquitous Secure Technology (TRUST) are also working on cybersecurity solutions in the broad Internet context but are not focusing on the issues faced by distributed science. NSF Middleware Initiative (NMI)-funded projects such as Shibboleth and PKI-Lite are addressing some of the authentication and authorization issues important to secure collaborative science. Although these programs are focused on cybersecurity tool development, the cybersecurity and operational needs of the distributed science community are not currently on the agenda of most of these programs.

The Department of Homeland Security is currently targeting cybersecurity solutions that can be deployed within the year and are aiming at community sites such as local fire stations that currently have little or no cybersecurity. They are also targeting deployments of corporate cybersecurity solutions to the various DHS agencies. Although distributed science is not directly a DHS concern, many of the mechanisms needed to protect distributed science are needed by other agencies.

5. Conclusion

The DOE Office of Science and the NSF provide many user facilities and are key participants in an increasing number of distributed collaborative science projects. Cyber attacks continue to increase in frequency and effectiveness — to date these attacks have not resulted in a major loss of data or capacity, but it is unlikely that this trend will continue. Cybersecurity mechanisms are needed which provide DOE Office of Science laboratories and NSF facilities adequate cybersecurity protection, while enabling an open science environment and allowing participation in global collaborations.

A Distributed Science Cybersecurity Program is needed to bring together a combination of COTS, leading research technologies, and research into the tools and other components needed to protect and maintain its operational environments; and to provide better coordination between security components and mechanisms to allow virtual organization job and resource monitoring are needed. Specific goals of the program include:

- *High Performance Computing and Networking:* This program will develop solutions for traffic analysis, intrusion detection, and intrusion prevention on

high-speed links. It will also create mechanisms to quickly detect and recover from intrusions on high performance computing resources.

- *Connecting Cybersecurity Components:* This program will develop a framework for coordination between security components and merging information from multiple security layers to provide a more complete view, providing for a more complete and accurate cybersecurity perspective.
- *Enabling and Protecting Virtual Organizations:* This program will develop tools and services to allow virtual organizations to better monitor their resources and perform incident containment.
- *Analysis/Response/Forensics/Recovery:* This program will develop tools for configuration verification and secure authenticated communication, as well as tools for aiding incident response. These are essential and often overlooked components to an overall security architecture.

A key to the program's success will be deployment and testing of cybersecurity infrastructure on real networks and systems, since it is impossible to replicate in a research network the scenarios experienced in a real network. Testing over real production networks is the only means of determining the utility and scalability of the approach, and is essential to the success of any security measure in this environment.

With this Distributed Science Cybersecurity Program, the DOE Office of Science and the NSF can be partners in developing a model for protecting open collaborative science.